Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12142–12149, October 1996
Biochemistry

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 25, 1995.

The N-end rule: Functions, mysteries, uses

(ubiquitin/proteolysis/N-degron/peptide import/apoptosis)

ALEXANDER VARSHAVSKY*

Division of Biology, California Institute of Technology, Pasadena, CA 91125

Contributed by Alexander Varshavsky, August 6, 1996

ABSTRACT **The N-end rule relates the *in vivo* half-life of a protein to the identity of its N-terminal residue. Similar but distinct versions of the N-end rule operate in all organisms examined, from mammals to fungi and bacteria. In eukaryotes, the N-end rule pathway is a part of the ubiquitin system. I discuss the mechanisms and functions of this pathway, and consider its applications.**

The half-lives of proteins in a living cell range from a few seconds to many days. Among the functions of intracellular proteolysis are the elimination of abnormal proteins, the maintenance of amino acid pools in cells affected by stresses such as starvation, and the generation of protein fragments that act as hormones, antigens, or other effectors. Yet another function of proteolytic pathways is selective destruction of proteins whose concentrations must vary with time and alterations in the state of a cell. Metabolic instability is a property of many regulatory proteins. A short *in vivo* half-life† of a regulator provides a way to generate its spatial gradients and allows for rapid adjustments of its concentration (or subunit composition) through changes in the rate of its synthesis. Conditionally unstable proteins, long-lived or short-lived depending on the state of a cell, are often deployed as components of control circuits. One example is cyclins—a family of proteins whose destruction at specific stages of the cell cycle regulates cell division and growth (3). In addition, many proteins are long-lived as components of larger complexes such as ribosomes and oligomeric proteins but are metabolically unstable as free subunits.

Features of proteins that confer metabolic instability are called degradation signals, or degrons (4). The essential component of one degradation signal, the first to be discovered, is a destabilizing N-terminal residue of a protein (5–7). This signal is called the N-degron. A set of N-degrons containing different destabilizing residues yields a rule, termed the N-end rule, which relates the *in vivo* half-life of a protein to the identity of its N-terminal residue (Table 1 and Fig. 1). The N-end rule pathway is present in all organisms examined, including the bacterium *E. coli* (8, 10), the yeast (fungus) *S. cerevisiae* (5, 9), and mammalian cells (1, 11) (Fig. 1).

The N-end rule was encountered in experiments that explored the metabolic fate of a fusion between ubiquitin (Ub) and a reporter protein such as *E. coli* β-galactosidase (βgal) in *S. cerevisiae* (5). In yeast and other eukaryotes, Ub-X-βgal is cleaved, cotranslationally or nearly so, by Ub-specific processing proteases at the Ub-βgal junction. This cleavage takes place regardless of the identity of the residue X at the C-terminal side of the cleavage site, proline being the single exception. By allowing a bypass of the normal N-terminal processing of a newly formed protein, this finding (Fig. 2*A*) yielded an *in vivo* method for generating different residues at the N termini of otherwise identical proteins—a technical advance that led to the N-end rule (5, 6).

In eukaryotes, the N-degron comprises at least two determinants: a destabilizing N-terminal residue and an internal lysine (or lysines) of a substrate (Fig. 2*B*) (9, 12, 14, 15). The Lys residue is the site of formation of a multi-Ub chain (16, 17). Ub is a 76-residue protein whose covalent conjugation to other proteins is involved in a multitude of processes—cell growth and differentiation, signal transduction, DNA repair, transmembrane traffic, and responses to stress, including the immune response. In many of these settings, Ub acts through routes that involve processive degradation of Ub-protein conjugates (18–21).

The binding of an N-end rule substrate by a targeting complex is followed by formation of a substrate-linked multi-Ub chain (22, 23). The ubiquitylated substrate is processively degraded by the 26S proteasome—an ATP-dependent, multisubunit protease (18, 20, 24, 25). The N-end rule pathway is present in both the cytosol (1, 5, 11) and the nucleus (J. A. Johnston and A.V., unpublished data). In this paper, I summarize the current understanding of the N-end rule. For a more detailed review, see ref. 26.

Definitions of Terms

The N-End Rule. A relation between the metabolic stability of a protein and the identity of its N-terminal residue.

The N-Degron. For a degradation signal to be termed an N-degron, it is necessary and sufficient that it contain a substrate's initial or acquired N-terminal residue whose recognition by the targeting machinery is essential for the activity of this degron.

The Pre-N-Degron. Features of a protein that are necessary and sufficient, in the context of a given intracellular compartment, for the formation of N-degron.

The N-End Rule Pathway. A set of molecular components that is necessary and sufficient, in the context of a given intracellular compartment, for the recognition and degradation of proteins bearing N-degrons. This "hardware-centric" definition of the pathway bypasses semantic problems that arise if, for example, one and the same targeting complex recognizes not only N-degrons but also a degradation signal whose essential determinants do not include the N-terminal residue of a substrate. This definition also encompasses a setting where N-degrons that bear different destabilizing N-terminal residues are recognized by distinct targeting complexes. As discussed below, neither of these possibilities is entirely hypothetical.

Primary Destabilizing Residues. Destabilizing activity of these N-terminal residues, denoted N-d^p, requires their physical binding by a protein called N-recognin or E3. In eukaryotes, the type

Abbreviations: Ub, ubiquitin; βgal, β-galactosidase; DHFR, dihydrofolate reductase; N-d^t, a tertiary destabilizing N-terminal residue; N-d^s, a secondary destabilizing N-terminal residue; N-d^{p1} and N-d^{p2}, type 1 and type 2 primary destabilizing N-terminal residues, respectively; Nt-amidase, N-terminal amidohydrolase; Nt^N-amidase, amidohydrolase specific for N-terminal Asn; R-transferase, Arg-tRNA-protein transferase; L/F-transferase, Leu/Phe-tRNA-protein transferase; ICE, interleukin-1β-converting enzyme; ER, endoplasmic reticulum; comtoxin, codominance-mediated toxin; NLS, nuclear localization signal.

*To whom reprint requests should be addressed. e-mail: varshavskya@starbase1.caltech.edu.

†The *in vivo* degradation of many short-lived proteins, including the engineered N-end rule substrates, deviates from first-order kinetics (1, 2). Therefore the term "half-life," if applied to an entire decay curve, is a useful but often crude approximation. A more rigorous terminology for describing nonexponential decay was proposed by Lévy *et al.* (1).

Table 1. The N-end rule in *E. coli* and *S. cerevisiae*

Residue X in X-βgal	*In vivo* half-life of X-βgal, min	
	In *E. coli*	In *S. cerevisiae*
Arg	2	2
Lys	2	3
Phe	2	3
Leu	2	3
Trp	2	3
Tyr	2	10
His	>600	3
Ile	>600	30
Asp	>600	3
Glu	>600	30
Asn	>600	3
Gln	>600	10
Cys	>600	>1200
Ala	>600	>1200
Ser	>600	>1200
Thr	>600	>1200
Gly	>600	>1200
Val	>600	>1200
Pro	?	?
Met	>600	>1200

Approximate *in vivo* half-lives of X-β-galactosidase (βgal) proteins in *E. coli* at 36°C (8) and in *S. cerevisiae* at 30°C (5, 9). A question mark at Pro indicates its uncertain status in the N-end rule (see text).

1 binding site of N-recognin binds N-terminal Arg, Lys, or His—a set of basic N-d^p residues, whereas the type 2 site binds N-terminal Phe, Leu, Trp, Tyr, or Ile—a set of bulky hydrophobic N-d^p residues (Figs. 3 and 5). Accordingly, the N-d^p residues are subdivided into type 1 (N-d^{p1}) and type 2 (N-d^{p2}) residues. The N-d^p residues of *E. coli*—Phe, Leu, Trp, and Tyr—are exclusively type 2 (N-d^{p2}) residues (Figs. 4 and 5*C*).

Secondary Destabilizing Residues. These N-terminal residues, denoted N-d^s, are Arg and Lys in *E. coli*; Asp and Glu in *S. cerevisiae*; and Asp, Glu, and Cys in mammalian cells (Figs. 3–5). In eukaryotes, destabilizing activity of N-d^s residues requires their accessibility to Arg-tRNA-protein transferase (R-transferase). In bacteria such as *E. coli*, destabilizing activity of the N-d^s residues Arg and Lys requires their accessibility to Leu/Phe-tRNA-protein transferase (L/F-transferase).

Tertiary Destabilizing Residues. N-terminal Asn and Gln residues, denoted N-d^t. Destabilizing activity of N-d^t residues requires their accessibility to N-terminal amidohydrolase (Nt-amidase) (Fig. 5 *A* and *B*).

Stabilizing Residues. A stabilizing N-terminal residue is a "default" residue, in that it is stabilizing because targeting components of an N-end rule pathway do not bind to it (or modify it) efficiently enough even in the presence of other determinants of an N-degron. Gly, Val, and Met are stabilizing residues that are common to all of the known N-end rules (Fig. 1).

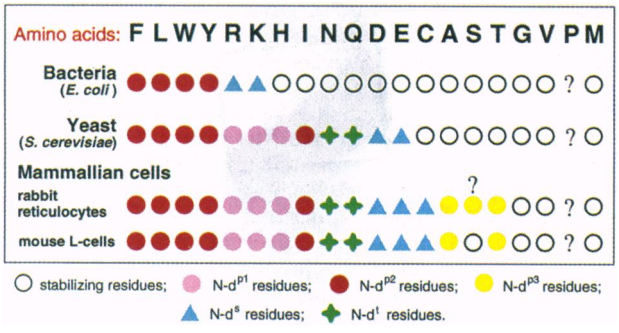

FIG. 1. Comparison of eukaryotic and bacterial N-end rules. Open circles denote stabilizing residues. Purple, red, and yellow circles denote, respectively, type 1, type 2, and type 3 primary destabilizing residues (N-d^{p1}, N-d^{p2}, and N-d^{p3}). Blue triangles denote secondary destabilizing residues (N-d^s). Green crosses denote tertiary destabilizing residues (N-d^t) (1, 5–11). A question mark at Pro indicates its uncertain status (see the main text). A question mark above Ser indicates its uncertain status in the reticulocyte N-end rule (1).

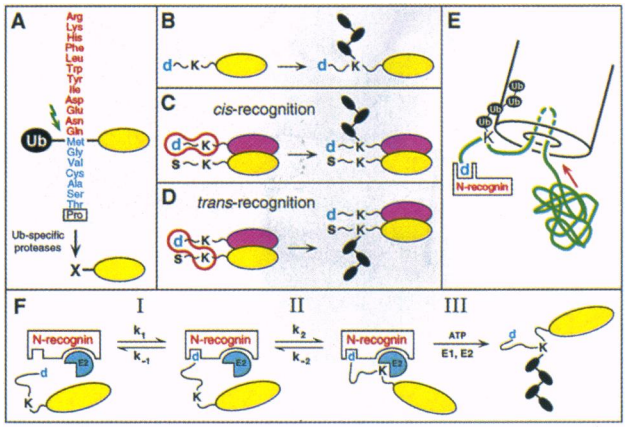

FIG. 2. Mechanics of the N-end rule. (*A*) The Ub fusion technique. Linear fusions of Ub to other proteins are cleaved at the last residue of Ub, making it possible to produce different residues at the N-termini of otherwise identical proteins (5, 11). Amino acid residues in blue and red are stabilizing and destabilizing, respectively, in the *S. cerevisiae* N-end rule (9). (*B*) The two-determinant organization of eukaryotic N-degrons. d, a destabilizing N-terminal residue. A chain of black ovals linked to the second-determinant lysine (K) denotes a multi-Ub chain. (*C*) Cis recognition of the N-degron in one subunit of a dimeric protein. The other subunit bears s, a stabilizing N-terminal residue. (*D*) Trans recognition, in which the first (d) and second (K) determinants of the N-degron reside in different subunits of a dimeric protein (12). (*E*) The hairpin insertion model. A targeted N-end rule substrate (in green) bearing a multi-Ub chain is shown bound to the 26S proteasome through the chain. The position of a targeting complex containing N-recognin is unknown, and is left unspecified. Only the 20S core component of the 26S proteasome is shown. A red arrow indicates the direction of net movement of the substrate's polypeptide chain toward active sites in the interior of proteasome. By analogy with the arrangement of signal sequences during transmembrane translocation of proteins (13), it is proposed that a region of the substrate upstream of its ubiquitylated lysine (K) does not move through the proteasome during the substrate's degradation, and it may be released intact following a cleavage in the vicinity of the lysine. Variants of this model may also be relevant to the targeting of proteins that bear internal or C-terminal degrons. (*F*) A model for the recognition of an N-end rule substrate (9). The reversible binding of N-recognin to a primary destabilizing N-terminal residue (d) of a substrate (step I) must be followed by a capture of the second-determinant lysine (K) of the substrate by a targeting complex containing a Ub-conjugating (E2) enzyme (step II). It is unknown whether the lysine is captured by E2 (as shown here) or by N-recognin. Ubiquitylation of the substrate commences once the targeting complex is bound to both determinants of the N-degron (step III). This model does not specify, among other things, the details of Ub conjugation (see the main text).

Components and Evolution of the N-End Rule Pathway

N-Recognin (E3). In *S. cerevisiae*, N-recognin is a 225-kDa protein (encoded by *UBR1*) that selects potential N-end rule substrates through the binding to their N-d^p residues Phe, Leu, Trp, Tyr, Ile, Arg, Lys, or His (26, 27). N-recognin has at least two substrate-binding sites. The type 1 site is specific for the basic N-terminal residues Arg, Lys, and His. The type 2 site is specific for the bulky hydrophobic N-terminal residues Phe, Leu, Trp, Tyr, and Ile (Fig. 3). At present, these sites are defined through dipeptide-based competition experiments. Specifically, a dipeptide bearing a destabilizing N-terminal residue was found to inhibit the degradation of a test N-end rule substrate if that substrate's N-terminal residue was of the same type (1 or 2) as the dipeptide's N-terminal residue (2, 11, 28).

A genetic dissection of the type 1 and type 2 sites in *S. cerevisiae* N-recognin (Ubr1p) has shown that either of the sites can be mutationally inactivated without significantly perturbing the other site. Mutations that selectively inactivate the type 1 or the type 2 site are located within the ≈50-kDa N-terminal region of the 225-kDa N-recognin (A. Webster, M. Ghislain, and A.V., unpublished data). E3α, the mammalian counter-

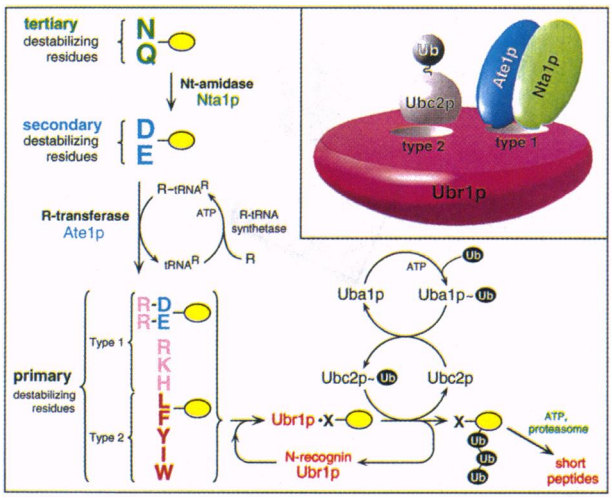

FIG. 3. The *S. cerevisiae* N-end rule pathway. Type 1 (N-d^{p1}) and type 2 (N-d^{p2}) primary destabilizing N-terminal residues are in purple and red, respectively. Secondary (N-d^{s}) and tertiary (N-d^{t}) destabilizing N-terminal residues are in blue and green, respectively. The yellow ovals denote the rest of a protein substrate. The conversion of N-d^{t} residues N and Q into N-d^{s} residues D and E is mediated by N-terminal amidohydrolase (Nt-amidase), encoded by *NTA1*. The conjugation of the N-d^{p1} residue R to N-d^{s} residues D and E is mediated by Arg-tRNA-protein transferase (R-transferase), encoded by *ATE1*. A complex of N-recognin and the Ub-conjugating (E2) enzyme Ubc2p catalyzes the conjugation of activated Ub, produced by the Ub-activating (E1) enzyme Uba1p, to a Lys residue of the substrate, yielding a substrate-linked multi-Ub chain. Uba1p~Ub and Ubc2p~Ub denote covalent (thioester-mediated) complexes of these enzymes with Ub. A multiubiquitylated substrate is degraded by the 26S proteasome. (*Inset*) A model of the targeting complex. The 20-kDa Ubc2p E2 enzyme is depicted carrying activated Ub linked to Cys-88 of Ubc2p through a thioester bond. Both the 52-kDa Nta1p (Nt-amidase) and the 58-kDa Ate1p (R-transferase) bind to the 225-kDa Ubr1p (N-recognin) in proximity to the type 1 substrate-binding site of Ubr1p. In addition, Nta1p directly interacts with Ate1p (see the main text).

part of *S. cerevisiae* N-recognin, has been characterized biochemically in extracts from rabbit reticulocytes (18). Another mammalian N-recognin, termed E3β, which apparently binds to substrates bearing N-terminal Ala and Thr (and possibly also Ser) (1, 11), has been described as well (18).

All eukaryotes examined have both Ub and the N-end rule pathway. Some, but not all, prokaryotes contain Ub (29). The bacterium *E. coli* lacks Ub but does have an N-end rule pathway (Fig. 4) (8). Screens for mutations that inactivate this

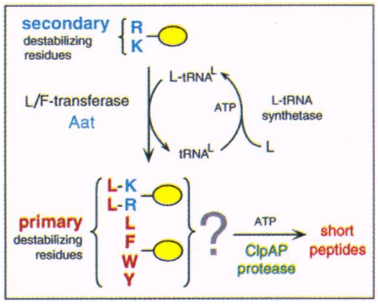

FIG. 4. The *E. coli* N-end rule pathway. Primary (N-d^{p}) destabilizing N-terminal residues L, F, W, and Y are in red. Secondary (N-d^{s}) destabilizing N-terminal residues R and K are in blue. The yellow ovals denote the rest of a protein substrate. Conjugation of the N-d^{p} residue L to the N-d^{s} residues R and K is mediated by Leu/Phe-tRNA-protein transferase (L/F-transferase), encoded by *aat* (8). *In vivo*, L/F-transferase appears to conjugate predominantly, if not exclusively, L (10). The degradation of a substrate bearing an N-d^{p} residue is carried out by the ATP-dependent protease ClpAP, encoded by *clpA* and *clpP*. A question mark denotes an ambiguity about the nature of N-recognin in *E. coli*.

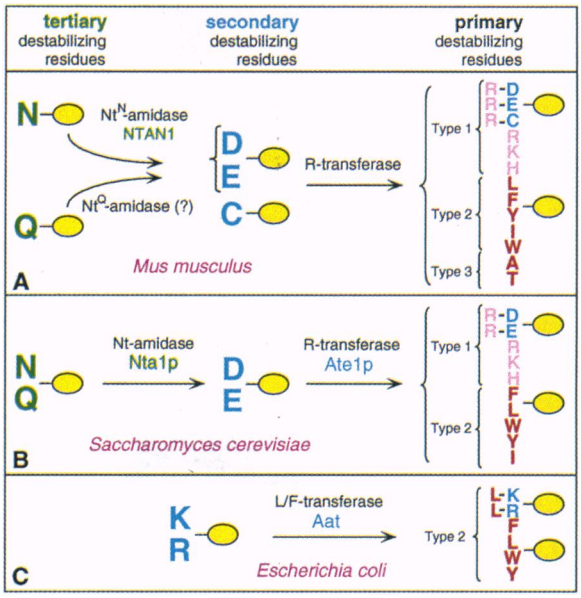

FIG. 5. Comparison of enzymatic reactions that underlie the activity of tertiary (N-d^{t}) and secondary (N-d^{s}) destabilizing residues in different organisms. (*A*) Mouse (*Mus musculus*) L-cells and rabbit (*Oryctolagus cuniculus*) reticulocytes (1, 11). (*B*) The yeast *S. cerevisiae* (9). (*C*) The bacterium *E. coli* (8, 10). The *E. coli* N-end rule lacks N-d^{t} residues. The postulated mammalian Nt^{Q}-amidase (in *A*) remains to be identified.

pathway have identified three *E. coli* genes—*clpA*, *clpP*, and *aat* (8, 10). Aat is a Leu/Phe-tRNA-protein transferase (L/F-transferase). ClpA (81 kDa) and ClpP (21 kDa) form an ≈750-kDa complex, ClpAP, which exhibits ATP-dependent protease activity *in vitro* (30) and is a functional counterpart of the eukaryotic 26S proteasome in the *E. coli* N-end rule pathway (Fig. 4). ClpP exhibits a chymotrypsin-like proteolytic activity *in vitro* (30). ClpA is the ATP-binding component of ClpAP. *In vitro* studies have shown that ClpA can act as a chaperone in the activation of RepA, the replication initiator encoded by the plasmid P1 (32). *In vivo* ramifications of these results, and in particular their relevance to the proteolytic function of ClpAP in the *E. coli* N-end rule pathway (Fig. 4), remain to be examined.

N-Terminal Amidases. The *S. cerevisiae* N-terminal amidohydrolase (Nt-amidase), encoded by *NTA1*, is a 52-kDa enzyme which deamidates Asn or Gln if they are located at the N-terminus of a polypeptide (Figs. 3 and 5*B*) (31, 33).

Stewart *et al.* (34) purified a porcine Nt-amidase that deamidates N-terminal Asn (N) but not Gln (Q), and isolated a cDNA that encodes this enzyme. Grigoryev *et al.* (33) isolated and characterized an ≈17-kb gene, termed *Ntan1*, that encodes a mouse homolog of the porcine amidase, termed Nt^{N}-amidase. The ≈1.4-kb *Ntan1* mRNA is expressed in all of the tested mouse tissues and cell lines. Construction of *ntan1*Δ mouse mutants is under way (Y. T. Kwon and A.V., unpublished data).

Aminoacyl-tRNA-Protein Transferases. The *S. cerevisiae* Arg-tRNA-protein transferase (R-transferase), encoded by *ATE1*, is a 58-kDa enzyme which utilizes Arg-tRNA to arginylate N-termini of polypeptides that bear Asp or Glu (Figs. 3 and 5*B*) (35). In contrast to *S. cerevisiae*, where only Asp and Glu are N-d^{s} residues, in mammals Cys is an N-d^{s} residue as well (11) (Table 1 and Fig. 1).

An extract prepared ≈2 hr after a crush injury to the rat sciatic nerve from a segment of the nerve immediately upstream of the crush site was found to conjugate an ≈10-fold higher amount of [^{3}H]arginine to the N termini of unidentified endogenous proteins than an otherwise identical extract from the same region of an unperturbed sciatic nerve (36). This finding suggested a crush-induced increase in the level of N-end rule substrates and/or a post-crush induction of the N-end rule pathway. No post-crush increase in arginylation was observed with extracts

from the optic nerve, which does not regenerate after a crush injury, in contrast to the sciatic nerve (36).

R-transferase appears to be confined to eukaryotes, whereas L/F-transferase is present in bacteria such as *E. coli* but is apparently absent from eukaryotes. *E. coli* L/F-transferase is a 27-kDa enzyme encoded by the *aat* gene (10). *In vivo*, L/F-transferase conjugates mainly, if not exclusively, Leu to N-terminal Arg or Lys of a polypeptide substrate (10) (Fig. 4). *E. coli* mutants lacking *aat* are unable to degrade N-end rule substrates that bear N-terminal Arg or Lys. These data (8, 10) identified L/F-transferase as a component of the *E. coli* N-end rule pathway.

Ub-Conjugating Enzymes. The initial interaction between an N-end rule substrate and N-recognin is of moderate affinity (the inferred $K_d \approx 10$ μM; ref. 26), but it becomes much stronger if an internal lysine of the substrate is captured by a targeting complex containing a Ub-conjugating (E2) enzyme and N-recognin (E3). This capture initiates processive synthesis of a lysine-linked multi-Ub chain. The E2 enzymes utilize activated Ub, produced by the Ub-activating (E1) enzyme, to catalyze the formation of isopeptide bonds between the C-terminal Gly-76 of Ub and ε-amino groups of lysines in acceptor proteins (Fig. 3) (18, 19, 37).

In at least some Ub-dependent systems (38), including apparently the N-end rule pathway (V. Chau and A.V., unpublished data), the pathway-specific Ub ligase—a complex of a recognin (E3) and an E2 enzyme—catalyzes the transfer of the Ub moiety (which is initially linked to a Cys residue of the E1 enzyme) through a relay of Ub thioesters before conjugating Ub to a Lys residue of a targeted substrate. In a substrate-linked multi-Ub chain, the C-terminal Gly of one Ub moiety is joined to an internal Lys of the adjacent Ub moiety, resulting in a chain of Ub-Ub conjugates. In a multi-Ub chain linked to an N-end rule substrate, only Lys-48 of Ub was found to be joined to another Ub moiety within a chain (16). Recently, multi-Ub chains linked through Lys-63, Lys-29, Lys-11, or Lys-6 of Ub have been described as well (16, 17, 39–41). It is unknown whether these chains play a role in the N-end rule pathway.

The N-End Rule As a Witness of Evolution. The hierarchic organization of N-end rules, with their tertiary, secondary, and primary destabilizing residues, is a feature more conserved in evolution than either the Ub dependence of an N-end rule pathway or the identity of enzymatic reactions that mediate the activity of destabilizing residues. For example, in a bacterium such as *E. coli*, which lacks the Ub system, the N-end rule has both N-d^s and N-d^p residues (it lacks N-d^t residues) (Figs. 1, 4, and 5*C*). The identities of N-d^s residues in *E. coli* (Arg and Lys) are different from those in eukaryotes (Figs. 1 and 5). Bacterial and eukaryotic enzymes that implement the coupling between N-d^s and N-d^p residues are also different: L/F-transferase in *E. coli* and R-transferase in eukaryotes. Note, however, that bacterial L/F-transferase and eukaryotic R-transferase catalyze reactions of the same type (conjugation of an amino acid to an N-terminal residue of a polypeptide) and utilize the same source of activated amino acid (aminoacyl-tRNA) (Fig. 5).

The apparent confinement of R-transferase to eukaryotes and of L/F-transferase to prokaryotes suggests that N-d^s residues were recruited late in the evolution of N-end rule, after the divergence of prokaryotic and eukaryotic lineages. The lack of sequence similarity between the yeast Nt-amidase and the mammalian Nt^N-amidase, as well as the more narrow specificity of the mammalian enzyme (Fig. 5 *A* and *B*) suggest that the N-d^t residues Asn and Gln became a part of the N-end rule much later yet, possibly after the divergence of metazoan and fungal lineages. If so, the N-end rule pathway may be an especially informative witness of evolution: the ancient origins of this proteolytic system, the simplicity and discreteness of changes in the rule books of N-end rules among different species, and the diversity of proteins that either produce or target the N-degron should facilitate phylogenetic deductions—once the components of this pathway become characterized across a broad range of organisms.

Code Versus Hardware. A given N-end rule is defined operationally—for a set of proteins such as X-βgals that differ exclusively by their N-terminal residues. Existing evidence (9) strongly suggests that the ranking aspect of an N-end rule—i.e., an ordering of relative destabilizing activities among 20 fundamental amino acids—is invariant from one protein reporter to another in a given intracellular compartment. By contrast, the actual *in vivo* half-lives may differ greatly among *different* proteins bearing one and the same N-terminal residue (9). The cause of these differences is the multicomponent nature of underlying N-degrons (Fig. 2*B*).

A priori, one and the same N-end rule can be implemented through vastly different assortments of targeting hardware. At one extreme, each destabilizing N-terminal residue may be bound by a distinct N-recognin. Conversely, a single N-recognin may be responsible for the entire rule book of destabilizing residues in a given N-end rule. The actual N-end rule pathways lie between these extremes and happen to have a hierarchic rather than "linear" structure (Figs. 3–5).

Targeting Complex of the N-End Rule Pathway

The known components of the *S. cerevisiae* N-end rule pathway that mediate steps prior to the processive proteolysis of a targeted substrate by the 26S proteasome are Nt-amidase (Nta1p), R-transferase (Ate1p), N-recognin (Ubr1p), a Ub-conjugating (E2) enzyme (Ubc2p), and the Ub-activating (E1) enzyme (Uba1p) (Fig. 3) (22, 26, 27, 31, 35, 42). In addition to a direct evidence for the physical association between N-recognin and Ubc2p (22, 23), there is also circumstantial evidence for the existence of a complex between N-recognin, R-transferase, and Nt-amidase (31). Recently, a high-affinity interaction between Nta1p and Ate1p was demonstrated directly; other evidence suggests that both Nta1p and Ate1p interact with Ubr1p (M. Ghislain, A. Webster, and A.V., unpublished results). In a quaternary Ubc2p–Ubr1p–Nta1p–Ate1p complex suggested by these data, Ate1p and Nta1p interact with each other and with Ubr1p (Fig. 3).

Effects of overexpressing Nt-amidase and/or R-transferase in *S. cerevisiae* not only suggested the existence of Nta1p–Ate1p–Ubr1p–Ubc2p complex but also led to the prediction that Nta1p and Ate1p are associated with Ubr1p in proximity to its type 1 substrate-binding site (31). In Fig. 3 diagram, a physical proximity of the bound R-transferase to the type 1 site of N-recognin is presumed to decrease the steric accessibility of this site to an N-end rule substrate that bears an N-d^{p1} residue such as Arg and approaches the type 1 binding site of N-recognin from the bulk solvent. By contrast, a substrate that acquired Arg through arginylation by the N-recognin-bound R-transferase would be able to reach the (nearby) type 1 binding site of N-recognin directly—without dissociating into the bulk solvent first—a feature known as substrate "channeling" in multistage enzymatic reactions (43). The mechanics of channeling may involve diffusion of an N-end rule substrate in proximity to surfaces of the targeting complex, analogous to the mechanism of a bifunctional enzyme dihydrofolate reductase (DHFR)–thymidylate synthetase, where the channeling of dihydrofolate apparently results from its movement across the surface of the protein (44).

The N-Degron and Pre-N-Degron

Nascent proteins contain N-terminal Met (fMet in prokaryotes), which is a stabilizing residue in the known N-end rules (Fig. 1). Thus, the N-degron of an N-end rule substrate must be produced from a pre-N-degron. In an engineered N-end rule substrate, a pre-N-degron contains the N-terminal Ub moiety whose removal by Ub-specific proteases yields the protein's N-degron (Fig. 2*A*). This design of a pre-N-degron is unlikely to be relevant to physiological N-end rule substrates, because natural Ub fusions (including the precursors of Ub) either contain a stabilizing residue at the Ub–protein junction or bear a mutant Ub moiety that is retained *in vivo* (45–47). The known Met-aminopeptidases remove N-terminal Met if the second residue of a protein is stabilizing in the yeast N-end rule (Fig. 1). The structural basis of this selectivity is the size

of a residue's side chain (48–50). Specifically, the side chains of the residues that are destabilizing in the yeast N-end rule are larger than those of stabilizing residues. The exception is Met—a bulky hydrophobic but stabilizing residue (Fig. 1).

Can there be just one or a few residues between N-terminal Met and the site of cleavage that produces an N-degron? If so, a short (≤10 residues) N-terminal sequence might contain both the recognition motif and the cleavage site(s) for a relevant (unknown) processing protease. Screens for such sequences, carried out in *S. cerevisiae* (51, 52), did identify short (≤10 residues) N-terminal regions that conferred Ubr1p-dependent metabolic instability on a reporter protein. Most of the sequences identified by these screens were not similar to each other, possibly because a very large number of 10-residue N-terminal extensions can produce an N-degron *in vivo*, analogous to a large number of N-terminal sequences that can function as signals for protein translocation across the endoplasmic reticulum (ER) membrane (53).

Analysis of one N-terminal extension identified by Ghislain *et al.* (52) has shown that it targets a reporter protein for degradation while retaining its N-terminal Met (M. Gonzalez, F. Lévy, M. Ghislain, and A.V., unpublished data). This finding suggests that N-recognin binds not only to N-degrons but also to a degron that consists of an entirely internal sequence motif. By contrast, two other examined extensions were found to be cleaved after N-terminal Met, yielding destabilizing N-terminal residues (51, 52). In sum, we are just beginning to understand the processing reactions that yield a destabilizing N-terminal residue in a non-polyprotein context.

Mechanics of N-Degron

Stochastic Capture Model. Studies with βgal- and DHFR-based N-end rule substrates (9, 12, 16) suggested a stochastic view of the N-degron, in which specific Lys residues of an N-end rule substrate can be assigned a probability of being utilized as a ubiquitylation site. This probability depends on time-averaged spatial position and mobility of a protein's Lys. For some, and often for most, of the Lys residues in an N-end rule substrate, the probability of serving as a ubiquitylation site would be negligible because of the Lys's lack of mobility and/or its distance from a destabilizing N-terminal residue. In this "stochastic capture" model (Fig. 2*F*), the folded conformation of a substrate would be expected to slow down or preclude the search for a Lys residue, unless it is optimally positioned in the folded substrate.

The bipartite design of N-degron (Fig. 2*B*) is likely to be also characteristic of other Ub-dependent degradation signals—present in a multitude of naturally short-lived proteins that include cyclins (3), IκBα (54), and c-Jun (21). The first component of these degrons is an internal region of a protein (instead of its N-terminal residue) that is specific for each degradation signal. The second component is an internal Lys residue (or residues). A degron may also contain regulatory determinants whose modification (e.g., phosphorylation/dephosphorylation) can modulate the activity of this degron (14, 21).

Cis–Trans Recognition and Subunit-Specific Degradation of Oligomeric Proteins. The two determinants of N-degron can be recognized either in cis or in trans (Fig. 2 *C* and *D*) (ref. 12; F. Lévy and A.V., unpublished data). Experiments that revealed the trans-recognition have also brought to light a remarkable feature of the N-end rule pathway: only those subunits of an oligomeric protein that contain the ubiquitylation site (but not necessarily a destabilizing N-terminal residue) are actually degraded (12). What might be the mechanism of subunit-specific proteolysis? A "simple" model is suggested by the binding of a substrate-linked multi-Ub chain to a component of the proteasome (55). Specifically, a subunit of an oligomeric substrate bound to the proteasome through a subunit-linked multi-Ub chain may be the only subunit that undergoes further mechanochemical processing by ATP-dependent, chaperone-like components of the 26S proteasome. These components mediate the unfolding and translocation steps that cause a movement of the subunit toward active sites in the proteasome's interior, and in the process dissociate this subunit from the rest of oligomeric substrate. In this mechanism, the initial binding of N-recognin to another subunit—the one that bears a destabilizing N-terminal residue but not the Lys determinant (Fig. 2*C*)—may be either too transient (lasting, in a "productive" engagement, only long enough for a Lys to be captured on a nearby subunit) or sterically unfavorable for the delivery of this subunit to the interior of the proteasome.

Since other Ub-dependent degradation signals appear to be organized similarly to the N-degron (a "primary" recognition determinant and an internal lysine or lysines), subunit selectivity is likely to be a general feature of proteolysis by the Ub system (6). Examples of physiologically relevant subunit-selective proteolysis include the degradation of p53 in a complex with the papilloma-viral protein E6 (38, 56) and the degradation of a cyclin in a complex with a cyclin-dependent kinase (3).

The Hairpin Insertion Model and the Function of the Multi-Ub Chain. Formation of a substrate-linked multi-Ub chain produces an additional binding site (or sites) for components of the proteasome (6, 55). The resulting increase in affinity—i.e., a decrease in the rate of dissociation of the proteasome–substrate complex— can be used to facilitate proteolysis. Suppose that a rate-limiting step which leads, several steps later, to the first proteolytic cleavage of the proteasome-bound substrate is an unfolding of a relevant region of the substrate. If so, an increase in stability of the proteasome–substrate complex, brought about by the multi-Ub chain, should facilitate substrate degradation, because the longer the allowed "waiting" time, the greater the probability of a required unfolding event. Another (not mutually exclusive) possibility is that a substrate-linked multi-Ub chain acts as a proximity trap for partially unfolded states of a substrate. This might be achieved through reversible interactions of the chain's Ub moieties with regions of the substrate that undergo local unfolding. A prediction common to both models is that the degradation of a substrate whose conformation poses less of a kinetic impediment to the proteasome should be less dependent on Ub and ubiquitylation than the degradation of an otherwise similar but more stably folded substrate.

How is a proteasome-bound, ubiquitylated protein directed to the interior of the proteasome? This problem is analogous to the one in studies of transmembrane channels for protein translocation (13). Could the solutions be similar in these systems, reflecting, perhaps, a common ancestry of translocation channels and proteasomes? A model in Fig. 2*E* proposes, by analogy with translocation systems, a "hairpin" insertion mechanism for the initiation of proteolysis by the 26S proteasome. A biased random walk ("thermal ratchet") that is likely to underlie the translocation of proteins across membranes (13) may also be responsible for the movement of the substrate's polypeptide chain through the proteasome, with cleavage products diffusing out from the proteasome's distal end and thereby contributing to the net bias in the chain's bidirectional saltations through the proteasome channel.

Two findings indicate that unfolding of a targeted N-end rule substrate is a prerequisite for its degradation by the 26S proteasome. Methotrexate—a folic acid analog and high-affinity ligand of DHFR—can inhibit the degradation of an N-end rule substrate such as Arg-DHFR by the N-end rule pathway (57). This result suggests that a critical post-ubiquitylation step faced by the proteasome includes a "sufficient" conformational perturbation of the proteasome-bound substrate. Further, it was shown that the N-end rule-mediated degradation of a 17-kDa N-terminal fragment of the 70-kDa Sindbis virus RNA polymerase is not precluded by the conversion of all of the fragment's 10 Lys residues into Arg residues, which cannot be ubiquitylated (T. Rümenapf, J. Strauss, and A.V., unpublished data). Thus, the ubiquitylation requirement of previously studied N-end rule substrates may be a consequence of their relatively stable conformations. The binding of a largely unfolded substrate (such as a fragment of Sindbis virus RNA polymerase) by the targeting complex of the N-end rule pathway may be sufficient for delivery of the substrate to the proteasome's active sites in the absence of a multi-Ub chain. In the language of models in Fig. 2 *E* and *F*,

the "waiting" time for a bound and conformationally unstable substrate may be short enough not to require the formation of a dissociation-slowing device such as a multi-Ub chain.

Substrates and Functions of the N-End Rule Pathway

The N-End Rule and Osmoregulation in Yeast. A synthetic lethal screen was used to isolate an *S. cerevisiae* mutant, termed *sln1* (for *s*ynthetic *l*ethal of *N*-end rule), whose viability requires the presence of *UBR1* (58). *SLN1* has been found to encode a eukaryotic homolog of two-component regulators—a large family of proteins previously encountered only in bacteria (59). The properties of *S. cerevisiae* Sln1p are consistent with it being a sensor component of the osmoregulatory (HOG) pathway (60). Since an otherwise lethal hypomorphic mutation in *SLN1* can be suppressed by the presence of Ubr1p (N-recognin) (58, 59), it is likely that one or more of the proteins (e.g., kinases) whose activity is down-regulated by Sln1p can also be down-regulated through their degradation by the N-end rule pathway. The relevant physiological N-end rule substrate(s) remains to be identified.

The N-End Rule and Peptide Import. Alagramam *et al.* (61) have found that *ubr1*Δ yeast cells are unable to import di- and tripeptides. Recent results (C. Byrd and A.V., unpublished data) indicate that Ubr1p (N-recognin) controls the activity of the peptide transporter Ptr2p, an integral plasma membrane protein, by regulating the synthesis and/or metabolic stability of *PTR2* mRNA. In one model, a transcriptional repressor of *PTR2* is short-lived, being degraded by the N-end rule pathway. Consistent with this mechanism, the control of *PTR2* expression by Ubr1p was found to involve the Ub-conjugating (E2) enzyme Ubc2p, a known component of the N-end rule pathway (Fig. 3). Ubc4p E2 enzyme can partially compensate for the absence of Ubc2p; a deletion of both *UBC2* and *UBC4* results in cells that do not express Ptr2p and are unable to import peptides, similarly to *ubr1*Δ cells.

A screen for mutants that allow a bypass of the requirement for *UBR1* in peptide import identified a gene, *CUP9*, that encodes a homeodomain-containing protein. Cup9p is short-lived; its degradation requires *UBR1* (C. Byrd and A.V., unpublished data). Cup9p is likely to be a transcriptional repressor of *PTR2*. Remarkably, an earlier study (62) identified *CUP9* as a gene whose inactivation decreases the resistance of *S. cerevisiae* to the toxicity of copper ions in the presence of a nonfermentable carbon source, suggesting a role for Cup9p in detoxification of copper. Although the connection between peptide import and resistance to copper toxicity remains obscure, our findings, taken together with the results of Knight *et al.* (62), suggest that the N-end rule pathway may be involved in the control of both peptide import and copper homeostasis.

On a Possible Function of the N-End Rule Pathway in Apoptosis. Until recently, apoptosis ("programmed" cell death) (63) was considered to be an attribute of multicellular but not unicellular organisms. However, given the near-identity of cells in a quasiclonal population of single-cell organisms, selection pressures may favor the emergence of an apoptotic response, for instance to a stress of starvation. By killing a fraction of a cell population, this response may benefit the rest of it. One example is the *mazEF* operon of *E. coli* that encodes a toxin/antitoxin pair (64). The long-lived MazF protein is toxic; the short-lived MazE binds to MazF and counteracts its toxicity. If the expression of the *mazEF* operon falls below a certain threshold, as can happen during starvation in *E. coli*, the level of antitoxin MazE would decrease more rapidly than the level of toxin MazF, resulting in a starvation-induced programmed cell death (64). Before the identification of *mazEF* in the *E. coli* chromosome, analogous pairs of genes ("addiction modules") have been found in a number of plasmids, where they ensure the plasmids' retention in their hosts. MazE is degraded by the protease ClpAP (64)—the same protease that degrades N-end rule substrates in *E. coli* (Fig. 4). It is unknown whether MazE contains a pre-N-degron or another degron recognized by ClpAP.

An essential aspect of apoptosis in metazoans may also be controlled by a pair (or pairs!) of proteins—"apoptosis modules"—that act similarly to bacterial addiction modules. Further, we suggest that the short-lived component of an apoptosis module may be an N-end rule substrate. One reason for considering this idea is the facility (a single cut) and irreversibility of a process that can convert an initially long-lived antitoxin component of an apoptosis module into a short-lived protein degraded by the N-end rule pathway. Specifically, we propose that the induction of apoptosis by interleukin-1β-converting enzyme (ICE)-like proteases (65, 66) may proceed through the cleavage of an (unknown) antitoxin component of an apoptosis module by an ICE or ICE-like protease. This cleavage, while not necessarily inactivating the antitoxin's function as such, would expose a destabilizing residue at the N terminus of a cleavage product, rendering it short-lived and thereby releasing a previously inhibited, relatively long-lived toxin. Implicit in this hypothesis is the assumption that a cleavage of antitoxin by an ICE-like protease is, by itself, insufficient (or not immediately sufficient) for the disruption of antitoxin's function, and that processive degradation of a C-terminal cleavage fragment of antitoxin by the N-end rule pathway is a required post-cleavage step.

The known targets of ICE-family proteases contain Asp at the P1 position and a small residue, typically Ala or Ser, at the P1′ position—the future N-terminus of the C-terminal cleavage fragment. Ala [and possibly also Ser in some settings (26)] is a weakly destabilizing N-d^p residue in the mammalian N-end rule (1, 11). However, ICE-family proteases can also cleave peptide bonds whose P1′ position is occupied by Asn—a strongly destabilizing N-d^t residue in the N-end rule (Fig. 5*A*). Examples of proteins cleaved by ICE-family proteases at the Asp-Asn bond include actin (65) and protein kinase Cδ (66). In addition, the activation of at least one ICE-family protease also involves its proteolytic cleavage at the Asp-Asn bond (67), suggesting that the enzymatic activation of this or analogous proteases may simultaneously render them short-lived *in vivo*—a possible source of negative control.

In sum, the hypothesis invokes an effector of apoptosis which is activated when its inhibitor is cleaved by an ICE-like protease, perhaps at the Asp-Asn bond, yielding a short-lived protein degraded by the N-end rule pathway. It is also possible that an N-end rule substrate regulates apoptosis upstream of ICE-like proteases. One prediction of the former model is that metabolic stabilization of the presumed Asn-bearing (cleaved) inhibitor, for example, through a perturbation of the N-end rule pathway, may inhibit the apoptosis. This prediction can be tested in mouse cells that lack the Asn-specific N-terminal amidase (Nt^N-amidase) (Fig. 5*A*) (33). Construction of *ntan1*Δ mouse mutants is under way (Y. T. Kwon and A.V., unpublished data).

Gα Subunit of G Protein. Overexpression of the N-end rule pathway was found to inhibit the growth of haploid but not diploid cells (68). This ploidy-dependent toxicity was traced to the enhanced degradation of Gpa1p, the Gα subunit of the G protein that regulates cell differentiation in response to mating pheromone. The half-life of newly formed Gα at 30°C is ≈50 min in wild-type cells, ≈10 min in cells overexpressing the N-end rule pathway, and >10 hr in cells lacking the pathway. The degradation of Gα is preceded by its multiubiquitylation (68). Like other Gα subunits of G proteins, the *S. cerevisiae* Gpa1p bears a conjugated N-terminal myristoyl moiety, which appears to be retained during the targeting of Gpa1p for degradation. A deletion of the first 88 residues of Gpa1p greatly accelerates its degradation but retains the requirement for Ubr1p (K. Madura, unpublished data). These data suggest that Ubr1p recognizes a feature of Gα that is distinct from the N-degron. Another, N-degron-based model invokes a trans-targeting mechanism (Fig. 2 *C* and *D*). A G_s-type Gα is short-lived in mouse cells as well (69), consistent with the possibility that Gα subunits of other organisms are also degraded by the N-end rule pathway. The activation of mouse Gα shortens its *in vivo* half-life (69), suggesting an adaptation-related function of Gα degradation.

Sindbis Virus RNA Polymerase and Other Viral Proteins. The Sindbis virus RNA polymerase, also called nsP4 (*n*on*s*tructural *p*rotein *4*), is produced by an endoproteolytic cleavage of the viral precursor polyprotein nsP1234 (70). The nsP4 protein bears N-terminal Tyr (an N-d^{P2} residue; Figs. 1 and 5*A*), and is degraded by the N-end rule pathway in reticulocyte extract (71). Tyr is an N-terminal residue of other alphaviral RNA polymerases as well (70), suggesting that these homologs of Sindbis virus RNA polymerase are also degraded by the N-end rule pathway. Whereas the bulk of newly formed nsP4 is rapidly degraded, a fraction of nsP4 in infected cells is long-lived, presumably within a replication complex that contains viral and host proteins (70).

There are many potential N-end rule substrates derived from viral polyproteins (72). One of them is the integrase of the human immunodeficiency virus (HIV), produced by cleavages within the gag-pol precursor polyprotein. The processed integrase bears N-terminal Phe (72), a strongly destabilizing N-d^{P2} residue in the N-end rule (Fig. 5*A*). Thus, it is possible that, similarly to the Sindbis virus RNA polymerase, at least a fraction of HIV integrase is short-lived *in vivo*.

c-Mos, a Protooncoprotein. This 39-kDa Ser/Thr-kinase is expressed predominantly in male and female germ cells. Sagata and colleagues (14, 73) have identified c-Mos as a physiological substrate of the N-end rule pathway that is targeted for degradation through its N-terminal Pro residue. Met-Pro-Ser-Pro, the encoded N-terminal sequence of *Xenopus* c-Mos, is conserved among all vertebrates examined (73). Since the N-terminal Met-Pro peptide bond is readily cleaved by the major cytosolic Met-aminopeptidases (48–50), the initially second-position Pro is expected to appear at the N terminus of nascent c-Mos cotranslationally or nearly so.

The activity of the Pro-based N-degron in c-Mos is inhibited through the phosphorylation of Ser-2 [Ser-3 in the c-Mos open reading frame (ORF)] (14, 73). During the maturation of *Xenopus* oocytes c-Mos is phosphorylated partially and reversibly, and therefore remains short-lived. Later—at the time of germinal vesicle breakdown and the arrest of mature oocytes (eggs) at the second meiotic metaphase, c-Mos becomes long-lived, owing to its nearly stoichiometric phosphorylation at Ser-2 (74). Fertilization or mechanical activation of a *Xenopus* egg releases the meiotic arrest through the induced degradation of c-Mos—caused by a nearly complete dephosphorylation of phosphoserine-2 (14, 73). Consistent with this model of the N-degron in c-Mos, the replacement of Ser-2 with Asp or Glu (whose negative charge mimics that of the phosphate group) rendered c-Mos long-lived, whereas the replacement of Ser-2 with Ala yielded a constitutively unstable c-Mos (73). Lys-33 (Lys-34 in the c-Mos ORF) is a major ubiquitylation site of the c-Mos N-degron (73).

In contrast to N-terminal Pro in the context of c-Mos, the N-terminal Pro followed by the sequence His-Gly-Ser-. . . [this is the context of engineered N-end rule substrates such as X-βgal and X-DHFR (5, 9)] did not confer a short half-life on a reporter protein in either yeast or mammalian cells (F. Lévy, T. Rümenapf, and A.V., unpublished data). One interpretation of these results is that the N-degron of c-Mos, whose conserved N-terminal sequence is Pro-Ser-Pro-. . . , has a "degron-enabling" internal determinant additional to, and perhaps specific for, the N-terminal Pro. The c-Mos N-degron is the first example of N-degron whose activity is regulated by phosphorylation (73).

Compartmentalized Proteins Retrotransported to the Cytosol. In contrast to cytosolic and nuclear proteins, the proteins that function in (or pass through) the ER, Golgi, and related compartments often bear destabilizing N-terminal residues—the consequence of cleavage specificity of signal peptidases, which remove signal sequences from proteins translocated into the ER (5, 6). Thus, one function of the N-end rule pathway might be the degradation of previously compartmentalized proteins that "leak" into the cytosol from compartments such as the ER (6). Remarkably, it has been found that at least some compartmentalized proteins can be retrotransported to the cytosol through a route that requires specific ER proteins. US11, the ER-resident transmembrane protein encoded by cytomegalovirus, causes the newly translocated major histocompatibility complex (MHC) class I heavy chain to be selectively retrotransported back to the cytosol, where the heavy chain is degraded by a proteasome-dependent pathway (75). Similarly, CPY*, a defective vacuolar carboxypeptidase of *S. cerevisiae*, is retrotransported to the cytosol shortly after entering the ER, and is degraded in the cytosol by a Ub/proteasome-dependent pathway that requires the Ubc7p Ub-conjugating enzyme (76). Whether the N-end rule pathway plays a role in the degradation of retrotransported proteins remains to be determined.

Applications of N-Degron

The portability and modular organization of N-degrons make possible a variety of applications whose common feature is the conferring of a constitutive or conditional metabolic instability on a protein of interest.

The N-Degron and Conditional Mutants. A frequent problem with conditional phenotypes is their leakiness—i.e., unacceptably high residual activity of either a temperature-sensitive (*ts*) protein at nonpermissive temperature or a gene of interest in the "off" state of its promoter. Another problem is the "phenotypic lag," which often occurs between the imposition of nonpermissive conditions and the emergence of a relevant null phenotype.

In one application of the N-end rule to the problem of phenotypic lag, a constitutive N-degron (produced as described in Fig. 2*A*) was fused to a protein expressed from an inducible promoter (77, 78). This method is constrained by the necessity of using a heterologous promoter and by a constitutively short half-life of a target protein, whose levels may therefore be suboptimal under permissive conditions. An alternative design is a portable, heat-inducible N-degron that is inactive at a low (permissive) temperature but becomes active at a high (nonpermissive) temperature (15). Linking this degron to proteins of interest yields a new class of *ts* mutants, called *td* (*t*emperature-activated *d*egron). The *td* method (15) does not require an often unsuccessful search for a *ts* mutation in a gene of interest.

The N-Degron and Conditional Toxins. A major limitation of the current pharmacological strategies stems from the absence of drugs that are specific for two or more independent molecular targets. For reasons discussed elsewhere (7, 79), it is desirable to have a therapeutic agent that requires the presence of two or more predetermined targets in a cell for it to be killed, and that would spare a cell if it lacks even one of these targets. Combining two "conventional" drugs against two different targets in a multidrug regimen would not attain this goal, because the two drugs together would perturb not only cells containing both targets but also cells containing just one of the targets. More generally, it is desirable to have drugs that exhibit a combinatorial selectivity, killing (or otherwise modifying) a cell if, and only if, it *contains* a predetermined set of molecular targets and at the same time *lacks* another predetermined set of molecular targets. Therapeutic agents of this, currently unrealistic, selectivity are likely to be free of side effects—the bane of present-day therapies against diseases such as cancer.

A strategy for designing reagents that are sensitive to the presence or absence of more than one target at the same time has recently been proposed (7, 79). The key feature of new reagents, termed *comtoxins* (*co*dominance-*m*ediated *toxins*), is their ability to utilize codominance, a property characteristic of many signals in proteins, including degrons and nuclear localization signals (NLSs). Codominance refers to the following property of these signals: each degron (or NLS) in a protein bearing two or more degrons (or NLSs) can target the protein for degradation (or transport to the nucleus) independently of other degrons (or NLSs) in the same protein. The crucial property of a degron-based comtoxin is that its intrinsic toxicity is the same in all cells, whereas its half-life (and, consequently, its steady-state level and overall toxicity) in a cell depends on the cell's protein composition, specifically on the presence of "target" proteins that have been chosen to define the profile of a cell to be eliminated. The target proteins would

bind to their ligands in a comtoxin molecule, and either physically obstruct the recognition of degrons or inactivate them catalytically, for example by phosphorylation. These and related ideas are described elsewhere (7, 79).

We are exploring the feasibility of comtoxins by using the N-degron as a degradation signal and the cytotoxic A-chains of ricin or diphtheria toxin as effector domains (T. Suzuki, I. V. Davydov, and A.V., unpublished data). Our current aim is to determine whether the concept of comtoxins can be implemented in the "easy" setting of a cell culture—without addressing, yet, the delivery problem, the immunogenicity of protein drugs, and related concerns.

Epilogue

Although many things have been learned about the N-end rule since its discovery 10 years ago, several key questions remain unanswered or glimpsed at best. For example, the detailed mechanics of targeting is not understood. Biochemical dissection of the N-end rule pathway reconstituted *in vitro* from defined (cloned) components will be essential for attaining this goal. Crystallographic-quality structural information about N-recognin and the entire targeting complex will be required as well. The recently emerged possibility that N-recognin may target not only N-degrons but also other degradation signals adds yet another level of complexity, which will have to be addressed.

Genetic screens for proteins degraded by the N-end rule pathway are our best hope for bringing to light physiological N-end rule substrates. It is already clear that at least some of these substrates are conditionally unstable—for example, partitioned between a short-lived free substrate and a long-lived complex of the substrate with other proteins. In addition, for some substrates, the rate-limiting step in their degradation may be a processing (cleavage) event that produces an N-degron from a pre-N-degron. If so, a significant fraction of extant substrate molecules may bear a stabilizing N-terminal residue. Given these obstacles to identifying physiological N-end rule substrates, they are likely to be more numerous than is apparent at the present time.

I am grateful to the current and former members of the laboratory whose work on the N-end rule is described in this review. I thank colleagues whose names are cited in the text for their permission to discuss unpublished data. Our studies are supported by grants from the National Institutes of Health and the Association for the Cure of Cancer of the Prostate.

1. Lévy, F., Johnsson, N., Rümenapf, T. & Varshavsky, A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 4907–4912.
2. Baker, R. T. & Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1090–1094.
3. Murray, A. & Hunt, T. (1993) in *The Cell Cycle* (Freeman, NY), pp. 60–62.
4. Varshavsky, A. (1991) *Cell* **64,** 13–15.
5. Bachmair, A., Finley, D. & Varshavsky, A. (1986) *Science* **234,** 179–186.
6. Varshavsky, A. (1992) *Cell* **69,** 725–735.
7. Varshavsky, A. (1996) *Cold Spring Harbor Symp. Quant. Biol.* **60,** 461–478.
8. Tobias, J. W., Shrader, T. E., Rocap, G. & Varshavsky, A. (1991) *Science* **254,** 1374–1377.
9. Bachmair, A. & Varshavsky, A. (1989) *Cell* **56,** 1019–1032.
10. Shrader, T. E., Tobias, J. W. & Varshavsky, A. (1993) *J. Bacteriol.* **175,** 4364–4374.
11. Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S. & Varshavsky, A. (1989) *J. Biol. Chem.* **264,** 16700–16712.
12. Johnson, E. S., Gonda, D. K. & Varshavsky, A. (1990) *Nature (London)* **346,** 287–291.
13. Schatz, G. & Dobberstein, B. (1996) *Science* **271,** 1519–1526.
14. Nishizawa, M., Furuno, N., Okazaki, K., Tanaka, H., Ogawa, Y. & Sagata, N. (1993) *EMBO J.* **12,** 4021–4027.
15. Dohmen, R. J., Wu, P. & Varshavsky, A. (1994) *Science* **263,** 1273–1276.
16. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D., Gonda, D. K. & Varshavsky, A. (1989) *Science* **243,** 1576–1583.
17. Spence, J., Sadis, S., Haas, A. L. & Finley, D. (1995) *Mol. Cell. Biol.* **15,** 1265–1273.
18. Hershko, A. (1991) *Trends Biochem. Sci.* **16,** 265–268.
19. Jentsch, S. (1992) *Annu. Rev. Genet.* **26,** 179–207.
20. Hochstrasser, M. (1995) *Curr. Biol.* **7,** 215–223.
21. Pahl, H. L. & Baeuerle, P. A. (1996) *Curr. Opin. Cell Biol.* **8,** 340–347.
22. Dohmen, R. J., Madura, K., Bartel, B. & Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7351–7355.
23. Madura, K., Dohmen, R. J. & Varshavsky, A. (1993) *J. Biol. Chem.* **268,** 12046–12054.
24. Rechsteiner, M., Hoffman, L. & Dubiel, W. (1993) *J. Biol. Chem.* **268,** 6065–6068.
25. Hilt, W. & Wolf, D. H. (1996) *Trends. Biochem. Sci.* **21,** 96–102.
26. Varshavsky, A., Byrd, C., Davydov, I. V., Dohmen, R. J., Du, F., *et al.* (1996) in *Ubiquitin and the Biology of the Cell*, eds. Finley, D. & Peters, J.-M. (Plenum, New York), in press.
27. Bartel, B., Wünning, I. & Varshavsky, A. (1990) *EMBO J.* **9,** 3179–3189.
28. Reiss, Y., Kaim, D. & Hershko, A. (1988) *J. Biol. Chem.* **263,** 2693–2698.
29. Wolf, S., Lottspeich, F. & Baumeister, W. (1993) *FEBS Lett.* **326,** 42–44.
30. Gottesman, S. & Maurizi, M. R. (1992) *Microbiol. Rev.* **56,** 592–621.
31. Baker, R. T. & Varshavsky, A. (1995) *J. Biol. Chem.* **270,** 12065–12074.
32. Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K. & Maurizi, M. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12218–12222.
33. Grigoryev, S., Stewart, A. E., Kwon, Y. T., Arfin, S. M., Bradshaw, R. A., Copeland, N. J. & Varshavsky, A. (1996) *J. Biol. Chem.* **271,** in press.
34. Stewart, A. E., Arfin, S. M. & Bradshaw, R. A. (1995) *J. Biol. Chem.* **270,** 25–28.
35. Balzi, E., Choder, M., Chen, W., Varshavsky, A. & Goffeau, A. (1990) *J. Biol. Chem.* **265,** 7464–7471.
36. Dayal, V. K., Chakraborty, G., Sturman, J. A. & Ingoglia, N. A. (1990) *Biochim. Biophys. Acta* **1038,** 172–177.
37. Pickart, C. (1988) in *Ubiquitin*, ed. Rechsteiner, M. (Plenum, New York), pp. 77–99.
38. Scheffner, M., Nuber, U. & Huibregtse, J. M. (1995) *Nature (London)* **373,** 81–83.
39. Arnason, T. A. & Ellison, M. J. (1994) *Mol. Cell. Biol.* **14,** 7876–7883.
40. Johnson, E. S., Ma, P. C. M., Ota, I. M. & Varshavsky, A. (1995) *J. Biol. Chem.* **270,** 17442–17456.
41. Baboshina, O. V. & Haas, A. L. (1996) *J. Biol. Chem.* **271,** 2832–2831.
42. McGrath, J. P., Jentsch, S. & Varshavsky, A. (1991) *EMBO J.* **10,** 227–237.
43. Negrutskii, B. S. & Deutscher, M. P. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4991–4995.
44. Knighton, D. R., Kan, C. C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M. & Matthews, D. A. (1994) *Nat. Struct. Biol.* **1,** 186–194.
45. Özkaynak, E., Finley, D., Solomon, M. J. & Varshavsky, A. (1987) *EMBO J.* **6,** 1429–1440.
46. Finley, D., Bartel, B. & Varshavsky, A. (1989) *Nature (London)* **338,** 394–401.
47. Watkins, J. F., Sung, P., Prakash, L. & Prakash, S. (1993) *Mol. Cell. Biol.* **13,** 7757–7765.
48. Li, X. & Chang, Y. H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 12357–12361.
49. Arfin, S. M. & Bradshaw, S. A. (1988) *Biochemistry* **27,** 7979–7984.
50. Sherman, F., Stewart, J. W. & Tsunasawa, S. (1985) *Bioessays* **3,** 27–31.
51. Sadis, S., Atienza, C. & Finley, D. (1995) *Mol. Cell. Biol.* **15,** 4086–4095.
52. Ghislain, M., Dohmen, R. J., Lévy, F. & Varshavsky, A. (1996) *EMBO J.*, in press.
53. Kaiser, C. A., Preuss, D., Grisafi, P. & Botstein, D. (1987) *Science* **235,** 312–317.
54. Scherer, D. C., Brockman, J. A., Chen, Z. J., Maniatis, T. & Ballard, D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 11259–11263.
55. Deveraux, Q., Ustrell, V., Pickart, C. & Rechsteiner, M. (1994) *J. Biol. Chem.* **269,** 7059–7061.
56. Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. (1990) *Cell* **63,** 1129–1136.
57. Johnston, J. A., Johnson, E. S., Waller, P. R. H. & Varshavsky, A. (1995) *J. Biol. Chem.* **270,** 8172–8178.
58. Ota, I. M. & Varshavsky, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2355–2359.
59. Ota, I. M. & Varshavsky, A. (1993) *Science* **262,** 566–569.
60. Maeda, T., Wurgler-Murphy, S. M. & Sato, H. (1994) *Nature (London)* **369,** 242–245.
61. Alagramam, K., Naider, F. & Becker, J. M. (1995) *Mol. Microbiol.* **15,** 225–234.
62. Knight, S. A. B., Tamai, K. T., Kosman, D. J. & Thiele, D. J. (1994) *Mol. Cell. Biol.* **14,** 7792–7804.
63. Ellis, R. E., Yuan, J. & Horvitz, H. R. (1991) *Annu. Rev. Cell Biol.* **7,** 663–698.
64. Aizenman, E., Engelberg-Kulka, H. & Glaser, G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 6059–6063.
65. Kayalar, C., Örd, T., Testa, M. P., Zhong, L.-T. & Bredesen, D. E. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2234–2238.
66. Emoto, Y., Manome, Y., Meinhardt, G., Kisaki, H., Kharbanda, S., Robertson, M., Ghayur, T., Wong, W. W., Kamen, R., Weichselbaum, R. & Kufe, D. (1995) *EMBO J.* **14,** 6148–6156.
67. Thornberry, N. A., Bull, H. G., Galaycay, J. R., Chapman, K. T., Howard, A. D., *et al.* (1992) *Nature (London)* **356,** 768–774.
68. Madura, K. & Varshavsky, A. (1994) *Science* **265,** 1454–1458.
69. Levis, M. J. & Bourne, H. R. (1992) *J. Cell Biol.* **119,** 1297–1307.
70. Strauss, J. H. & Strauss, E. G. (1994) *Microbiol. Rev.* **58,** 491–562.
71. deGroot, R. J., Rümenapf, T., Kuhn, R. J., Strauss, E. G. & Strauss, J. H. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 8967–8971.
72. Dougherty, W. G. & Semler, B. L. (1993) *Microbiol. Rev.* **57,** 781–822.
73. Nishizawa, M., Okazaki, K., Furuno, N., Watanabe, N. & Sagata, N. (1992) *EMBO J.* **11,** 2433–2446.
74. Watanabe, N., Hunt, T., Ikawa, Y. & Sagata, N. (1991) *Nature (London)* **352,** 247–248.
75. Wiertz, E. J. H. J., Jones, T. R., Sun, L., Bogyo, M., Geuze, H. J. & Ploegh, H. L. (1996) *Cell* **84,** 769–779.
76. Hiller, M. M., Finger, A., Scheiger, M. & Wolf, D. H. (1996) *Science* **273,** 1725–1728.
77. Park, E. C., Finley, D. & Szostak, J. W. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1249–1252.
78. Cormack, B. P. & Struhl, K. (1992) *Cell* **69,** 685–694.
79. Varshavsky, A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3663–3667.